AD-A034 479 WISCONSIN UNIV MADISON MATHEMATICS RESEARCH CENTER F/G 12/2
SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS BY MEANS OF FINITE E--ETC(U)
DEC 76 J B ROSSER DAAG29-75-C-0024
UNCLASSIFIED MRC-TSR-1705 NL

1 OF 1
AD A034479

END

DATE
FILMED
2 - 77

MRC Technical Summary Report #1705

SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS BY MEANS OF FINITE ELEMENTS. AN INTRODUCTORY SKETCH

J. Barkley Rosser

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

December 1976

Received May 6, 1976

Approved for public release
Distribution unlimited

Sponsored by

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS BY MEANS OF FINITE ELEMENTS AN INTRODUCTORY SKETCH

J. Barkley Rosser

Technical Summary Report #1705
December 1976

ABSTRACT

An introductory exposition is given covering the solution of some partial differential equations by means of the method of finite elements. Special attention is given to the means of getting numerical approximations to the answer.

AMS MOS Subject Classification 35A40 , 65N30

Key Words: Partial Differential Equations
Finite Element Methods

Work Unit No. 7 Numerical Analysis

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

SOLUTION OF PARTIAL DIFFERENTIAL EQUATIONS BY MEANS OF FINITE ELEMENTS

AN INTRODUCTORY SKETCH

J. Barkley Rosser

1. The mathematical framework. Engineers have long found it advisable to analyse the stresses and strains in structures that they design or construct. Not uncommonly, these structures are composed of many parts. By appealing to such principles as virtual work, equations can be written relating the stresses and strains of one part to those of adjacent parts. From this there results a set of simultaneous linear equations, which can be solved to get a complete picture of all the stresses and strains. For a structure with very many parts, the number of simultaneous equations will be very great. However, since most parts are adjacent to only a few other parts, most coefficients are zero. Hence, by ordering the equations judiciously, solutions can be obtained fairly quickly. Because the parts are separate and finite elements, the procedure became known as the method of finite elements. A short list of references to books on this subject is given at the end of the report. The most comprehensive one is that by Zienkiewicz, but others may be more suitable for inexperienced readers.

Sponsored by the United States Army under Contract No. DAAG29-75-C-0024.

The method had such success that it was adapted to many situations in which there were finite elements only by arbitrary definition. One such area is the approximate numerical solution of differential equations. We shall give an introductory sketch of the method as applied to certain partial differential equations. Specifically, consider the equation

$$(1.1) \qquad -\frac{\partial}{\partial x}\left(p\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(q\frac{\partial u}{\partial y}\right) + ru = f,$$

where u, f, p, q, and r are functions of x and y. We interpret L as that operator on u, depending on p, q, and r, such that Lu is the left side of (1.1). The specific problem is to find a u which satisfies (1.1) inside a given region Ω and satisfies some given conditions on the boundary $\Gamma(\Omega)$ of Ω. For example, we could ask that u satisfy (1.1) inside the square $0 < x < 1$, $0 < y < 1$, and ask further that $u = 0$ on the sides of the square.

We are not here seeking great generality. We shall consider regions Ω which are fairly simple, and have well behaved boundaries. Also, we shall ask that each of

$$\frac{\partial^2 u}{\partial x^2}, \quad \frac{\partial^2 u}{\partial y^2}, \quad \frac{\partial p}{\partial x}, \quad \frac{\partial q}{\partial y}, \quad r, \quad f$$

be continuous. This certainly assures that the left side of (1.1) makes sense; indeed is continuous.

We define the inner product $(u, v)_2$ by

$$(1.2) \qquad (u, v)_2 = \iint_{\Omega} uv \, dx \, dy .$$

Theorem 1.1. The equation (1.1) holds in Ω if and only if

$$(1.3) \qquad (Lu-f, v)_2 = 0$$

for all v which are continuous on Ω and equal to 0 on $\Gamma(\Omega)$.

Proof. If (1.1) holds, then obviously (1.3) holds. Now assume that (1.3) holds as stated. If (1.3) should hold for $v = Lu-f$, giving

$$(1.4) \qquad \iint_{\Omega} (Lu-f)^2 dx \, dy = 0,$$

then obviously $Lu - f$ would have to be zero all over Ω; that is, (1.1) holds. However, $Lu - f$ is likely not zero on $\Gamma(\Omega)$. What we do is take $v = Lu - f$ except very close to the boundary of Ω, and then bend it to be 0 at the boundary. Then the left sides of (1.3) and (1.4) will be nearly equal. By restricting the region where we bend $Lu - f$ to be closer and closer to the boundary, we make the left sides of (1.3) and (1.4) as nearly equal as we like. So (1.4) must hold.

We define the operator $M(u, v)$ by

$$(1.5) \qquad M(u, v) = \iint_{\Omega} \{p \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + q \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} + ruv\} dx \, dy.$$

Theorem 1.2. If v is 0 on $\Gamma(\Omega)$ and $\partial v/\partial x$ and $\partial v/\partial y$ are continuous, then

$$M(u, v) = (Lu, v)_2 . \tag{1.6}$$

Proof. In texts on advanced calculus, Green's theorem is stated in such forms as

$$\int_{\Gamma(\Omega)} (A\,dx + B\,dy) = \iint_{\Omega} \left(\frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}\right) dx\,dy. \tag{1.7}$$

Take

$$A = -vq\,\frac{\partial u}{\partial y}$$

$$B = vp\,\frac{\partial u}{\partial x} .$$

Then the left side of (1.7) is zero, because $v = 0$ on $\Gamma(\Omega)$. So we get

$$\iint_{\Omega} \left\{p\,\frac{\partial u}{\partial x}\,\frac{\partial v}{\partial x} + q\,\frac{\partial u}{\partial y}\,\frac{\partial v}{\partial y}\right\} dx\,dy = -\iint_{\Omega} v\left\{\frac{\partial}{\partial x}\left(p\,\frac{\partial u}{\partial x}\right) + \frac{\partial}{\partial y}\left(q\,\frac{\partial u}{\partial y}\right)\right\} dx\,dy.$$

From this, (1.6) easily follows.

We say that u is a generalized solution of our problem if it has continuous first derivatives and satisfies the given conditions on $\Gamma(\Omega)$ and if

$$M(u, v) = (f, v)_2 \tag{1.8}$$

for every v which is 0 on $\Gamma(\Omega)$ and has continuous first derivatives.

If u is a generalized solution, then by Theorem 1.2, we have $(Lu - f, v)_2 = 0$ for every v which is 0 on $\Gamma(\Omega)$ and has continuous first derivatives. However, we cannot use Theorem 1.1 directly to conclude that u satisfies (1.1), because Theorem 1.1 requires that (1.3) hold for still more v's. However, unless we are dealing with some very poorly behaved functions, it will usually be the case that if u is a generalized solution, it will indeed satisfy (1.1).

By reasoning as in the proof of Theorem 1.1, we conclude that if u is a generalized solution, then (1.8) holds for every v with continuous first derivatives.

Theorem 1.3. If p, q, and r are non-negative, then if u is a generalized solution, it is a w with continuous first derivatives taking the given values on $\Gamma(\Omega)$ that minimizes

$$M(w, w) - 2(f, w)_2. \tag{1.9}$$

Proof. Let w minimize (1.9). Take v with continuous first derivatives and equal to 0 on $\Gamma(\Omega)$. Then for real α, $w + \alpha v$ will have continuous first derivatives and take the given values on $\Gamma(\Omega)$. So

$$M(w + \alpha v, w + \alpha v) - 2(f, w + \alpha v)_2$$

must have a minimum at $\alpha = 0$. So its derivative must be zero at $\alpha = 0$. Expanding it gives

$$M(w, w) + 2\alpha M(w, v) + \alpha^2 M(v, v) - 2(f, w)_2 - 2\alpha(f, v)_2 .$$

Differentiating with respect to α and setting $\alpha = 0$ gives

$$M(w, v) = (f, v)_2 .$$

As this holds for all suitable v, we conclude that w is a generalized solution. Alternatively, let u be a generalized solution, and let v be another function with continuous first derivatives. Consider

$$Q = \{M(v, v) - 2(f, v)_2\} - \{M(u, u) - 2(f, u)_2\} .$$

Taking $v = u$ in (1.8) gives

$$Q = M(v, v) + M(u, u) - 2(f, v)_2 .$$

Using (1.8) in this gives

$$\begin{aligned} Q &= M(v, v) + M(u, u) - 2M(u, v) \\ &= M(u-v, u-v) \\ &= \iint_{\Omega} \{p \left(\frac{\partial(u-v)}{\partial x}\right)^2 + q \left(\frac{\partial(u-v)}{\partial y}\right)^2 + r(u-v)^2\} dx\, dy. \end{aligned}$$

As p, q, and r are non-negative, we get $Q \geq 0$. So u does produce a minimum.

With slightly stronger conditions on p, q, and r, we easily conclude that u is the unique w that minimizes (1.9).

This leads to the Rayleigh-Ritz procedure for approximating a generalized solution. Let $u_0, u_1, \ldots, u_n$ all have continuous first derivatives. Let u_0 take the given values on $\Gamma(\Omega)$. Let $u_1, u_2, \ldots, u_n$ all be zero on $\Gamma(\Omega)$. Let $c_1, c_2, \ldots, c_n$ be real parameters. Then

$$u_0 + \sum_{k=1}^{n} c_k u_k \tag{1.10}$$

has continuous first derivatives and takes the given values on $\Gamma(\Omega)$. We substitute (1.10) for w in (1.9), and choose the c's so as to minimize the resulting expression. To make this a minimum, the partial derivatives with respect to the various c_k's must all be zero. Taking the partial with respect to c_j gives

$$\sum_{k=1}^{n} c_k \iint_{\Omega} \{p \frac{\partial u_j}{\partial x} \frac{\partial u_k}{\partial x} + q \frac{\partial u_j}{\partial y} \frac{\partial u_k}{\partial y} + r u_j u_k\} dx\, dy$$

$$+ \iint_{\Omega} \{p \frac{\partial u_0}{\partial x} \frac{\partial u_j}{\partial x} + q \frac{\partial u_0}{\partial y} \frac{\partial u_j}{\partial y} + r u_0 u_j\} dx\, dy \tag{1.11}$$

$$= (f, u_j)_2 .$$

This is a set of simultaneous linear equations to be solved for the c's. What this does is to pick out the best approximation (in some sense) of the form (1.10) for the u in question. If we have chosen the u_k's cleverly, so that there is a good approximation to u of the form (1.10), we will have found it.

The Galerkin approach proceeds according to a different principle. By Theorem 1.1, we seek a u that satisfies (1.3) for all continuous v equal to 0 on $\Gamma(\Omega)$. If we can find a u that satisfies (1.3) for a judiciously chosen set of v_j, it cannot be too different from a u that satisfies (1.3) for all v. So try a u of the form (1.10) and let us require that it satisfy (1.3) for v successively taken equal to $v_1, v_2, \ldots, v_n$. For v_j, there results

$$\sum_{k=1}^{n} c_k \left(-\frac{\partial}{\partial x}\left(p \frac{\partial u_k}{\partial x}\right) - \frac{\partial}{\partial y}\left(q \frac{\partial u_k}{\partial y}\right) + r u_k,\ v_j\right)_2$$

$$+ \left(-\frac{\partial}{\partial x}\left(p \frac{\partial u_0}{\partial x}\right) - \frac{\partial}{\partial y}\left(q \frac{\partial u_0}{\partial y}\right) + r u_0,\ v_j\right)_2 \tag{1.12}$$

$$= (f,\ v_j)_2.$$

Again we have a set of simultaneous linear equations for the c_k. The resulting value of (1.10) not only satisfies (1.3) for each v_j, but for any linear combination of the v_j. If the v_j's are chosen so that arbitrary functions can be approximated well by linear combinations of the v_j's, we can hope that we have come close to satisfying (1.3) for arbitrary v's.

One can try a different attack by the Galerkin approach. The function u is a generalized solution if it satisfies (1.8) for arbitrary v. Let us try to satisfy (1.8) for v taken successively equal to $v_1, v_2, \ldots, v_n$. For v_j there results

$$\sum_{k=1}^{n} c_k \iint_{\Omega} \{p \frac{\partial v_j}{\partial x} \frac{\partial u_k}{\partial x} + q \frac{\partial v_j}{\partial y} \frac{\partial u_k}{\partial y} + r v_j u_k\} dx\, dy$$

$$+ \iint_{\Omega} \{p \frac{\partial u_0}{\partial x} \frac{\partial v_j}{\partial x} + q \frac{\partial u_0}{\partial y} \frac{\partial v_j}{\partial y} + r u_0 v_j\} dx\, dy \tag{1.13}$$

$$= (f, v_j)_2 .$$

It will be noted that if we take the v's the same as the u's, this reduces to (1.11), which was obtained by the Rayleigh-Ritz approach. However, this derivation of (1.11) is probably preferable because in deriving (1.11) by the Rayleigh-Ritz approach we worked from (1.9), in which w was prescribed to take certain values on $\Gamma(\Omega)$. For a u which is prescribed to satisfy mixed conditions on $\Gamma(\Omega)$ the derivation from (1.8) by the Galerkin approach is satisfactory.

An advantage of (1.13) over (1.12) is that for (1.12) one would have to choose the u's to be twice differentiable, whereas for (1.13) singly differentiable u's will suffice. This will be an important consideration later on.

2. Finite elements to the rescue. We now face up to the real problem, namely the judicious choice of the u's (and perhaps the v's) for the Rayleigh-Ritz or Galerkin approaches. It would be helpful if we could choose them so that the integrations inherent in (1.11), (1.12), and (1.13) could be easily performed. This is not absolutely necessary,

since we could grind out approximations on a computer by established quadrature formulas. It would be even more helpful if we could arrange that a large fraction of the coefficients of the c's would be zero. This would make the equations much easier to solve. It will turn out that this can be done.

To fix our ideas, let Ω be the square $0 < x < 1$, $0 < y < 1$; see Figure 1. We subdivide this into triangles, whose vertices form a grid

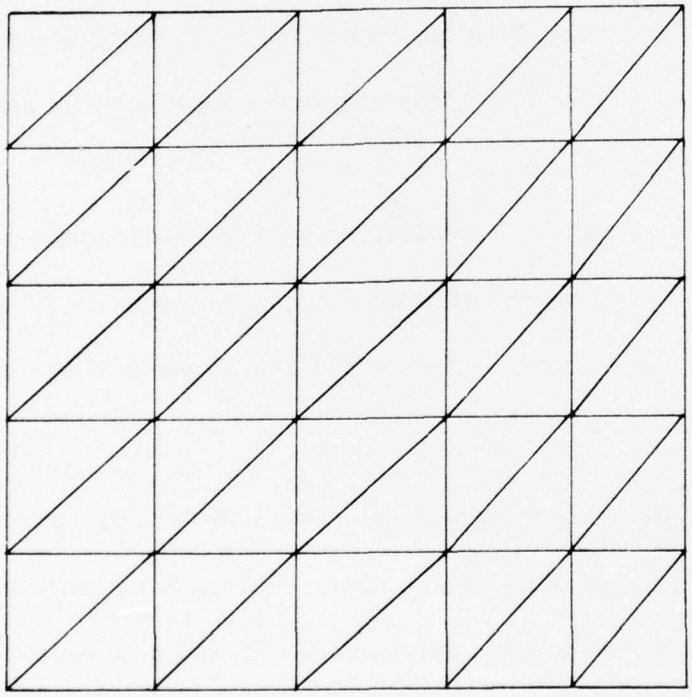

Figure 1.

on Ω. With each grid point we associate an area of support and a u_k. The area of support consists of all the triangles that have this grid point as a vertex. Then for each grid point we define the function u_k as follows. Outside the area of support, u_k is identically zero. Inside the area of support, u_k is a function which is 1 at the grid point and 0 along the sides of the triangles opposite the grid point. One very easy way to do this is to have u_k consist entirely of planes joined together along the sides of the triangles. Outside the area of support, u_k is a horizontal plane at zero altitude. For each triangle in the area of support, u_k is the plane which is one unit high at the grid point and zero units high along the opposite side of the triangle. That is, for the area of support u_k consists of a pyramid of height unity at the grid point, dropping down to zero at the edge of the area of support, while u_k is identically zero outside the area of support.

If now we try to approximate a function u which has the value c_k at the k-th grid point,

$$\sum_{k=1}^{n} c_k u_k \tag{2.1}$$

is obviously an approximating function. It has exactly the right value at each grid point. Not only that, but if the function u is fairly smooth, (2.1) is not a bad approximation anywhere. To be specific, let u have

continuous derivatives up to the second order, and let M_2 be the maximum absolute value of any derivative, first or second, in Ω. Then the difference between u and (2.1) is at most $4M_2h^2$, where h is the longest side of any triangle; see the Lemma on p. 142 of Prenter.

For any triangle in the area of support, it is easy to write the equation of the plane that constitutes u_k above that triangle. Let (x_0, y_0), (x_1, y_1), and (x_2, y_2) be the coordinates of the vertices of the triangle. We seek the plane which is one unit high at (x_0, y_0) and zero units high at each of (x_1, y_1) and (x_2, y_2). Inside the triangle we would have

$$(2.2) \qquad u_k = \frac{(x_1-x_2)(y_1-y_2)-(x_1-x_2)(y-y_2)+(x-x_1)(y_1-y_2)}{(x_1-x_2)(y_1-y_2)-(x_1-x_2)(y_0-y_2)+(x_0-x_1)(y_1-y_2)} .$$

The denominator is twice the area of the triangle, and so is not zero.

In (1.8) we specified that v was to have continuous first derivatives. But no u_k has continuous first derivatives. The discontinuities are not bad, just along the sides of certain triangles. Elsewhere, each u_k has continuous derivatives. It turns out that this is good enough. So we can take the v_j's in (1.12) or (1.13) to be the u_k's as defined above. Note that if (1.12) or (1.13) holds for all the v's, it holds for any linear combination of them. So in affect we are taking v to be (2.1). But, as we saw, any v can be represented fairly closely by (2.1). So we are close to assuring that (1.12) or (1.13) holds for any v.

It seems not unreasonable to try (2.1) as an approximation for u. If we do this, use of (1.12) will not be practical, since it requires second derivatives of u, whereas for the u_k defined above even the first derivatives leave something to be desired. However, we can use (1.13) perfectly well. So we use (1.13), taking the u_k's and v_j's to be just the u_k's defined above. This makes (1.13) the same as (1.11), except that more flexibility is permitted for the behavior of u along $\Gamma(\Omega)$. If the values of u are assigned along $\Gamma(\Omega)$, then we know the value to take for c_k at each boundary grid point. Thus we would take u_0 to be the sum of u_k's for boundary grid points multiplied by the respective boundary values.

In view of (2.2), the integrations in (1.13) can be carried out with the greatest of ease. Also, since each u_k is identically zero outside the area of support, many of the coefficients of the c_k in (1.13) will be zero; if we have a large number of mesh points most of the coefficients will be zero.

If we knew the values of u at each grid point, we could take the c_k to be these values, and then (2.1) would differ by less than $4M_2h^2$ from u at any point. However, the c_k are got by solving the equations (1.13). The resulting c_k will likely not be exactly the values of u at the k-th grid point. Indeed, there might be some question if they are even close. For the special case $p \equiv 1$, $q \equiv 1$, and $r \equiv 0$,

Prenter on pp. 231-236 presents a proof derived from Friedrichs that the c_k are close to the values of u at the grid points, so that (2.1) is close to u. In fact (see the top of p. 236 of Prenter), there is a constant N, independent of how the grid is constructed, so that

$$\iint_{\Omega} (u-u^*)^2 dx\, dy < \frac{Nh^2}{\sin^2\theta} \quad , \tag{2.3}$$

where u^* stands for (2.1), h is the longest side of a triangle, and θ is the smallest angle of a triangle. Hence, by going to a fine enough mesh, and being careful not to use triangles with small angles, one can contrive to make (2.1) as close as desired to the solution. Incidentally, in the third displayed formula on p. 233 of Prenter, one should put "K" for "$\sqrt{b-a}$".

As noted, the derivation in Prenter assumes $p \equiv 1$, $q \equiv 1$, and $r \equiv 0$. These requirements can be relaxed without invalidating the result. However, some restraints will still be needed. For example, in Theorem 1.3, the condition that p, q, and r are all non-negative was invoked; if one wished to conclude that u is the UNIQUE minimizing function, slightly more was required. Of course, we are working from (1.13), which did not depend for its motivation upon Theorem 1.3. However, if NO restraints are imposed on p, q, and r, there may fail to be a solution for (1.1). In any case, if the left side of (1.1) is to make any sense, some sort of differentiability seems to be required of p and q. The matter

seems to hinge upon whether the differential operator L is "strongly coercive." The short discussion on p. 254 of Prenter seems to indicate that if p and q have continuous first partial derivatives and are positive and bounded over Ω, and r is continuous, non-negative and bounded over Ω, then one may expect to get good approximations for u by solving (1.13). Also, the errors will be less than those indicated above (see p. 251 of Prenter). If p, q, and r fail to satisfy the conditions stated, advice from an expert should be sought. Incidentally, on p. 254 of Prenter, read "Borman" for "Birman" and change the last five words of the stated theorem to read: "... then A is strongly coercive."

It seems not agreed whether in this approach the "finite elements" are the triangles, or the u_k based on the grid points, or something else. Never mind, this is called a method of finite elements. It has many variations. For one thing, it can be used for solving ordinary differential equations; see pp. 201-227 of Prenter.

The results above can be generalized in various ways. The region Ω may be subdivided into rectangles; see pp. 116-137 of Prenter. Or one may take the u_k to consist of curved surfaces. A simple way to do this (discussed on pp. 144-153 of Prenter) proceeds as follows. In the triangular grid that we had before, introduce new grid points by taking one in each side of a triangle; usually the midpoint of the side is taken, but this is not obligatory. As before, each grid point is to have an area of support and a

u_k. For vertices of triangles, the area of support is as before. For each triangle in the area of support, u_k is given as the product of two formulas like the right side of (2.2) (see p. 147 of Prenter) and so is a quadratic polynomial. For grid points on the sides of triangles, the area of support is the pair of adjacent triangles (only one triangle if the grid point is on the boundary). For each triangle in the area of support, u_k is given as the product of two formulas like the right side of (2.2) (see p. 147 of Prenter). Outside the area of support, u_k is identically zero in all cases. At the grid point itself, u_k is unity, and it drops off to zero at the edge of the area of support; indeed this u_k is zero at all other grid points. Then (2.1) again serves as an approximating function. If u has continuous derivatives up to the third order, and M_3 is a bound of the absolute values of the derivatives, then (see p. 150 of Prenter) the difference between u and (2.1) is at most

$$\frac{8M_3h^3}{\sin\theta} ,$$

where h is the longest side of a triangle and θ is the least angle of a triangle.

Although these u_k's are curved surfaces, and make (2.1) fit u better than before, they still have discontinuous derivatives along sides of the triangles. Still more complicated u_k's have been devised, which have continuous derivatives everywhere. This increases the

difficulty of calculation quite a bit, and it is not at all clear that the gain is worth the extra effort. Many, many ingenious ways to define the u_k's have been proposed. On pp. 154-155 of Prenter are cited more than a dozen papers with proposed definitions.

3. Why use a finite element method? The equations, like (1.13), to be solved for the c_k's are similar to the equations which are to be solved in the finite difference approach. Also, the c_k's that are obtained are only approximations for u at the grid points. So the results, namely approximations to u at a set of grid points, are similar to those obtained by the finite difference approach. Nor does the finite element approach seem to produce appreciably smaller errors (or appreciably larger, either). However, there do seem to be two situations in which the finite element method appears preferable.

The first is the case where Ω has curved boundaries. Thus, let Ω be a circle, as in Figure 2. For the 28 triangles that border on the circle, we have the mismatch between the straight sides of the triangles and the circle. However, these triangles come fairly close to matching the circle. It certainly would not give a very bad approximation if we should solve (1.13) with the grid points shown in Figure 2.

A possible improvement would be to cut each of the 28 triangles that border on the circle in half by a line from the inward vertex to the circle. This would give us 28 more grid points, but each would be on the

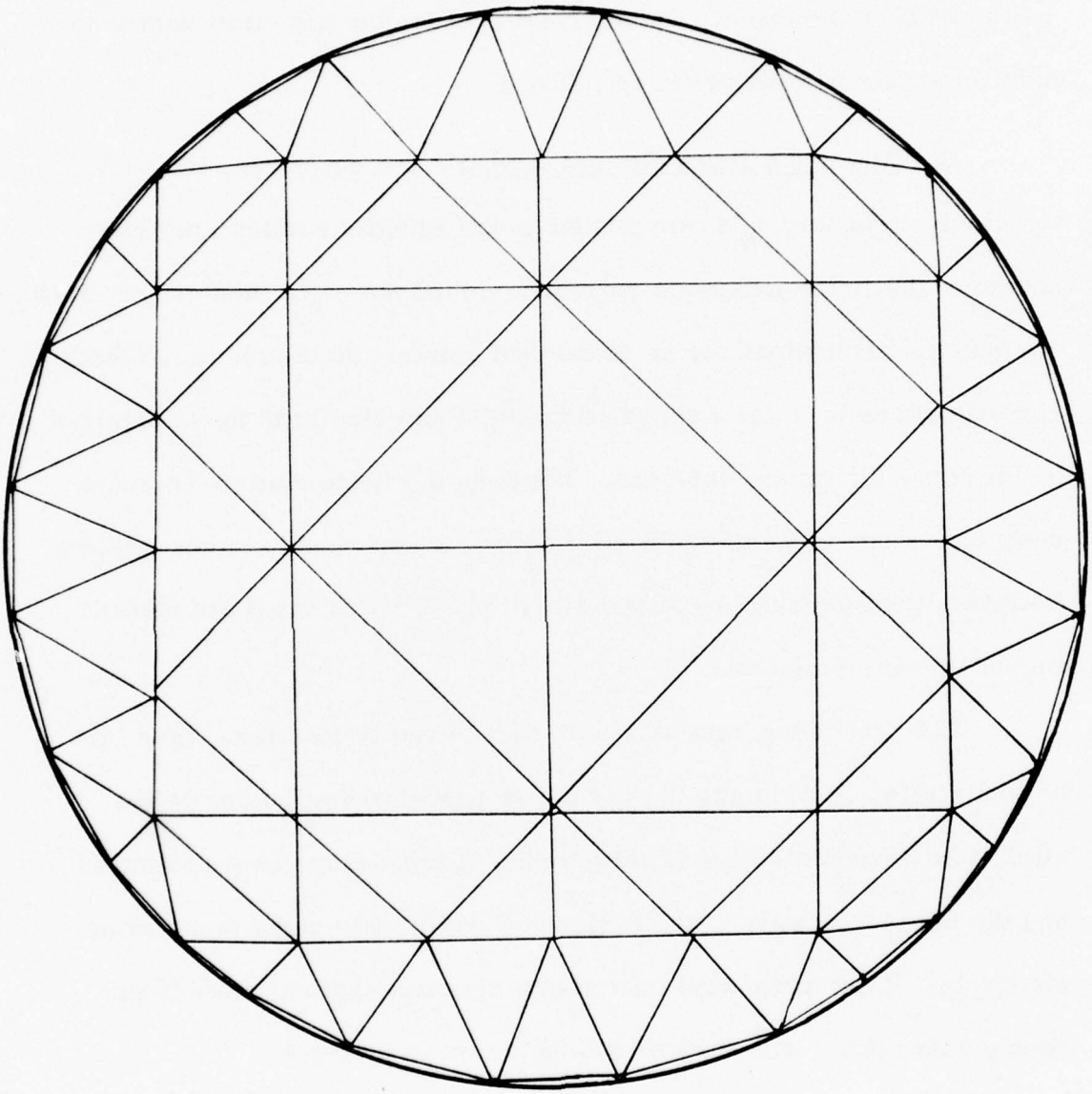

Figure 2.

circle. If values for u are assigned on the boundary, these extra grid points would cause no difficulty at all. We would now have 56 triangles bordering on the circle, which would match it much better. However, this would get us into smaller angles for the triangles, which would not be good in view of the $\sin^2\theta$ in the denominator on the right of (2.3).

Another way would be to add also new grid points at the mid-points of sides of triangles that do not border the circle. The triangles

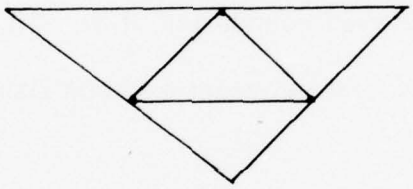

Figure 3.

that do not border the circle would be dissected as shown

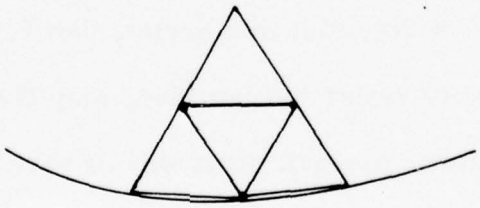

Figure 4.

in Figure 3, while those that border on the circle would be replaced by the configuration shown in Figure 4. This would approximately halve the h

in (2.3) while maintaining the θ about the same. Of course one does this at the cost of approximately doubling the number of equations to be solved. However, we now have 56 triangles bordering on the circle, which will certainly give a good fit. The problem of trying to handle this by finite differences seems to involve more extensive calculations.

A very sketchy discussion is given on pp. 155-167 of Prenter of the situation with curved boundaries. References are given to some papers in which Ω is divided into pieces which have some curved boundaries; see the remarks about "isoparametric transformations." This enables one to get better fits along the curved boundaries of Ω. All in all, if Ω has curved boundaries, use of finite element methods may well be quite advantageous.

When one has a reentrant angle in $\Gamma(\Omega)$, for instance if Ω is an L-shaped region, the first derivative of u is likely to be infinite in the neighborhood of the reentrant corner. A discussion of this situation is given in "Calculation of Potential in a Sector, Part I," by J. Barkley Rosser, MRC-Technical Summary Report Number 1535, May 1975. As a result, the usual finite difference approximation will be very poor indeed near the corner. While the finite difference equations can be solved in such a case, the resulting values for grid points near the reentrant corner will be very poor approximations. If one uses a finite element method, with a uniform

mesh, the approximation for grid points near the reentrant corner will be poor, and for similar reasons. However, it is very easy to refine the mesh near the reentrant corner; see Figure 5. With the smaller h near the corner, one will

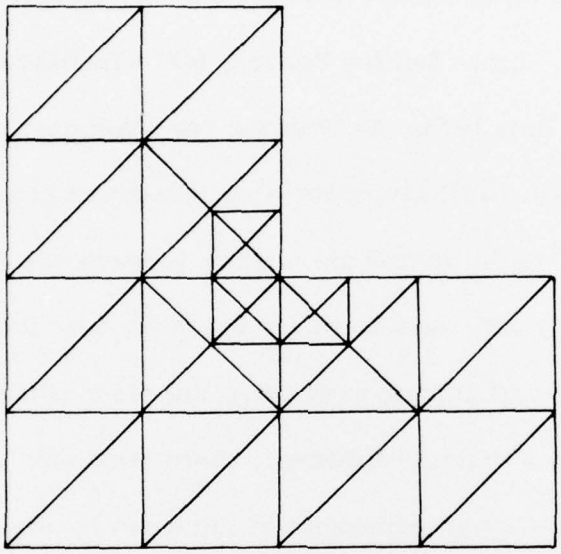

Figure 5

compensate in part for the larger derivatives. The local mesh refinement makes no trouble at all in the finite element method. It is possible to have a local mesh refinement with the finite difference method, but it involves

awkward interpolations, and the resulting equations to be solved do not fit into any of the schemes, such as S.O.R. or A.D.I., for abridging the calculations.

If one has discontinuous values assigned for u along the boundary, there will be infinite derivatives of u in the neighborhood of the discontinuities; see "Effect of Discontinuous Boundary Conditions on Finite-difference Solutions," by J. Barkley Rosser, MRC-Technical Summary Report Number 1383, June 1975. As with the reentrant corner, the finite difference approximation will give poor approximations near the discontinuities. The same would be true for a finite element solution of uniform mesh. However, it is very easy to refine the mesh near the discontinuities.

In fact, one might consider removing the discontinuity by the methods given on pp. 221-222 of Milne. However, there are cases (such as the reentrant corner) where local refinement of the mesh is very advantageous; for such cases, use of the finite element method is worthwhile.

4. A numerical example. Let us approximate the solution of

$$-\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2} = 2 \tag{4.1}$$

inside the unit square $0 < x < 1$, $0 < y < 1$, subject to the conditions that u shall be zero around the sides of the square.

It is easily verified that a solution is given by

$$(4.2)\quad u = x(1-x) - \sum_{m=0}^{\infty} \frac{8}{\pi^3(2m+1)^3} \frac{\sinh(2m+1)\pi y + \sinh(2m+1)\pi(1-y)}{\sinh(2m+1)\pi} \sin(2m+1)\pi x.$$

From this, high accuracy solutions will be calculated for comparison with the approximate solutions.

We superpose on the square a grid with the grid points 1/8 unit apart vertically and horizontally. We subdivide the square into triangles as in Figure 1. We then proceed to calculate the coefficients appearing in (1.13). We could calculate the partial derivatives of the u_k by (2.2), but there is a quicker way. We note that by (2.2) $\partial u_k/\partial x$ is constant over each triangle, and the same for $\partial u_k/\partial y$. So if we can determine the values of these at one point in a triangle, we know the values throughout the triangle. In Figure 6 we show a typical area of support for a grid point, G, at the center of the figure. Along FG, the value of u_k rises from 0 to unity in a distance of 1/8 unit. So $\partial u_k/\partial x = 8$, which holds

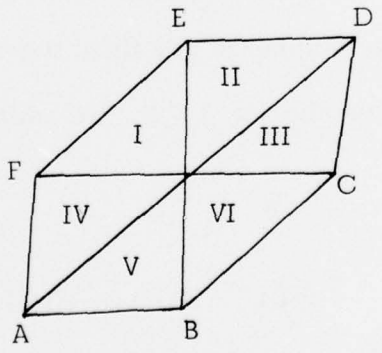

Figure 6.

throughout both the triangles I and IV. Along the line from G to C, $\partial u_k/\partial x = -8$, which value holds throughout both the triangles III and VI. Along ED and AB, we have $\partial u_k/\partial x = 0$, which value holds throughout both the triangles II and V. In a similar manner we find the value of $\partial u_k/\partial y$ to be: 8 throughout both the triangles V and VI; -8 throughout both the triangles I and II; 0 throughout both the triangles III and IV. Thus we readily calculate the products on the left of (1.13) inside each triangle. Integration is accomplished by multiplying by the area of the triangle.

Let us denote the grid point at $(\frac{i}{8}, \frac{j}{8})$ by G_{ij}, its associated u_k by u_{ij}, and the corresponding multiplier by c_{ij}. Since u is required to be 0 on the boundary, everything involving u_0 will be 0. So the left side of (1.13) reduces to

$$4c_{ij} - c_{(i-1)j} - c_{(i+1)j} - c_{i(j-1)} - c_{i(j+1)} .$$

On the right side of (1.13), the integration is performed by taking the volume of u_k over the area of support. But u_k is a pyramid over the area of support, whose volume is one third the base, 3/64, times the altitude, 1. As we have chosen f = 2, we multiply this by 2. So (1.13) takes the form

$$4c_{ij} - c_{(i-1)j} - c_{(i+1)j} - c_{i(j-1)} - c_{i(j+1)} = \frac{1}{32} \tag{4.3}$$

for each interior grid point G_{ij}.

It will be observed that these are exactly the same equations that would result from the finite difference method.

Although we have 49 equations in 49 unknowns, one can greatly simplify the calculations by appealing to symmetry. The solution is obviously symmetric about each diagonal of the square, and about the vertical and horizontal lines through the center. Consider the case of (4.3) that results when $i = j = 1$. Since u is zero around the boundary, we have $c_{01} = c_{10} = 0$. Also, by symmetry, we have $c_{21} = c_{12}$. So we get

$$4c_{11} - 2c_{21} = \frac{1}{32} . \tag{4.4}$$

Proceeding similarly, we get

$$4c_{21} - c_{11} - c_{31} - c_{22} = \frac{1}{32} \tag{4.5}$$

$$4c_{31} - c_{21} - c_{41} - c_{32} = \frac{1}{32} \tag{4.6}$$

$$4c_{41} - 2c_{31} - c_{42} = \frac{1}{32} \tag{4.7}$$

$$4c_{22} - 2c_{32} - 2c_{21} = \frac{1}{32} \tag{4.8}$$

$$4c_{32} - c_{22} - c_{42} - c_{31} - c_{33} = \frac{1}{32} \tag{4.9}$$

$$4c_{42} - 2c_{32} - c_{41} - c_{43} = \frac{1}{32} \tag{4.10}$$

$$4c_{33} - 2c_{43} - 2c_{32} = \frac{1}{32} \tag{4.11}$$

$$4c_{43} - 2c_{33} - c_{42} - c_{44} = \frac{1}{32} \tag{4.12}$$

$$4c_{44} - 4c_{43} = \frac{1}{32} . \tag{4.13}$$

We get (4.7) by using $c_{31} = c_{51}$, which follows by symmetry. We get (4.13) by using $c_{43} = c_{45} = c_{34} = c_{54}$, which follows by symmetry.

These ten equations are still quite sparse, so that their solution is quite easy; it can be carried out by a hand calculation. Note that (4.6) gives c_{31} in terms of three other c's. Now c_{31} occurs in only three other equations, so that one can easily eliminate c_{31}. We proceed to eliminate c_{11}, c_{31}, c_{22}, c_{42}, c_{33}, and c_{44} by means of equations (4.4), (4.6), (4.8), (4.10), (4.11), and (4.13). This is an instance of using a consistent ordering of the mesh points; see Smith, pp. 85-86.

There results

$$11c_{21} - c_{41} - 3c_{32} = \frac{7}{32} \tag{4.14}$$

$$9c_{32} - 3c_{21} - 2c_{41} - 3c_{43} = \frac{1}{4} \tag{4.15}$$

$$13c_{41} - 2c_{21} - 4c_{32} - c_{43} = \frac{7}{32} \tag{4.16}$$

$$7c_{43} - 6c_{32} - c_{41} = \frac{1}{4}\,. \tag{4.17}$$

We use (4.17) to eliminate c_{41} from the other three equations, after which the result of (4.16) can be used to eliminate c_{21}. That brings us down to two equations in two unknowns, and we get finally:

$$c_{11} = \frac{619}{17408} \cong 0.035558$$

$$c_{21} = \frac{966}{17408} \cong 0.055492$$

$$c_{31} = \frac{1146}{17408} \cong 0.065832$$

$$c_{41} = \frac{1202}{17408} \cong 0.069049.$$

From these, we get c_{22}, c_{32}, and c_{42} by (4.5), (4.6), and (4.7). Then we get c_{33} and c_{43} by (4.9) and (4.10). Finally we get c_{44} by (4.13). Approximate values are listed in Table 1.

We can improve our estimates by using Richardson's deferred approach to the limit; see Smith, pp. 140-141. This is not only valid for the finite difference method, but also for the finite element method because in this case the error also varies as h^2; see Prenter, p. 253. For this, we take a grid with 1/4 unit spacing between the grid points. This gives

$$c_{22} = \frac{11}{128}$$

$$c_{42} = \frac{14}{128}$$

$$c_{44} = \frac{18}{128} .$$

If we extrapolate from these, we get the values listed in Table 1.

TOP ENTRY IS ACCURATE
SECOND ENTRY FROM (4.4) TO (4.13)
THIRD ENTRY IS RICHARDSON EXTRAPOLATION
FOURTH ENTRY BASED ON (4.21)

			0.1473
			0.1456
			0.1472
			0.1405
		0.1321	0.1394
		0.1305	0.1377
		0.1323	0.1394
		0.1287	0.1353
	0.0906	0.1089	0.1147
	0.0893	0.1075	0.1133
	0.0905	0.1089	0.1146
	0.0886	0.1065	0.1120
0.0364	0.0564	0.0667	0.0699
0.0356	0.0555	0.0658	0.0690
0.0382	0.0567	0.0669	0.0699
0.0354	0.0551	0.0653	0.0685

Table 1.

From the more accurate extrapolated values, we can get better values for c_{11}, c_{31}, and c_{33}. If we rotate the square 45^{o}, equation (4.1) is invariant. So take a grid consisting of c_{11}, c_{31}, c_{33}, c_{22}, c_{42}, and c_{44}, and their reflections about the lines of symmetry. By finite differences, we will get

$$4c_{11} - c_{22} = \frac{1}{16} \tag{4.18}$$

$$4c_{31} - c_{22} - c_{42} = \frac{1}{16} \tag{4.19}$$

$$4c_{33} - c_{22} - 2c_{42} - c_{44} = \frac{1}{16} . \tag{4.20}$$

If we construct triangles on the same grid points, we will also get (4.18), (4.19), and (4.20) by finite elements. The improved values of c_{11}, c_{31}, and c_{33} from these equations have been entered in Table 1.

Finally, we use (4.5), (4.7), (4.9), and (4.12) to get improved values for c_{21}, c_{41}, c_{32}, and c_{43}; these have been entered in Table 1.

From Table 1, we see that the improved values are never off by more than 2 units in the third decimal, and mostly off by no more than 3 units in the fourth decimal. Considering that this resulted from a hand calculation, the agreement is very good indeed. Of course, we were able to take advantage of unusual symmetry. One cannot count on doing as well usually.

If one wishes greater accuracy, the obvious approach is to take a finer mesh. However, this considerably increases the computational labor. If a very fine mesh is taken, one can abridge the computations by appealing to S.O.R. or A.D.I. Since the equations are the same for finite differences and finite elements, these techniques are available for either.

Alternatively, one can use a more accurate approximating difference formula. See p. 143 of Smith, or the two reports: J. Barkley Rosser, "Nine-point Difference Solutions for Poisson's Equation," MRC-Technical Summary Report Number 1523, March 1975; and J. Barkley Rosser, "Finite-difference Solution of Poisson's Equation in Rectangles of Arbitrary Shape," MRC-Technical Summary Report Number 1404, February 1974. There has been some work on developing finite element approaches so as to decrease the error without refining the mesh, but the theory is not yet well advanced.

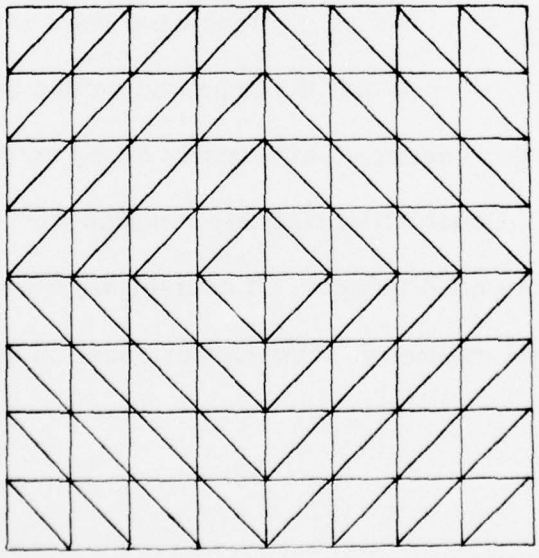

Figure 7.

One could triangulate the square as shown in Figure 7. This would produce the same set of equations as before, except that (4.13) would be replaced by

$$4c_{44} - 4c_{43} = \frac{1}{48} \quad . \tag{4.21}$$

For the solution, we get

$$c_{11} = \frac{462}{13056}$$

$$c_{21} = \frac{720}{13056}$$

$$c_{31} = \frac{853}{13056}$$

$$c_{41} = \frac{894}{13056} \quad .$$

From these, the other c's can be easily computed, as before. The results are listed in Table 1.

These results are uniformly lower than those based on the other triangulation, which was already uniformly low. If the reader is dismayed that two quite similar triangulations could give results as disparate as these, let him recall that these are only methods for obtaining APPROXIMATE values of u.

REFERENCES

GENERAL

Franz L. Alt and Morris Rubinoff, Eds., "Advances in Computers," Academic Press, 1962.

R. Courant and D. Hilbert, "Methods of Mathematical Physics," Interscience Publishers, 1953.

George E. Forsythe and Wolfgang R. Wasow, "Finite-difference Methods for Partial Differential Equations," John Wiley and Sons, 1960.

L. Fox, Ed., "Numerical Solution of Ordinary and Partial Differential Equations," Addison-Wesley Publishing Company, 1962.

W. E. Milne, "Numerical Solution of Differential Equations," John Wiley and Sons, 1960.

Richard S. Varga, "Matrix Iterative Analysis," Prentice-Hall, 1962.

FINITE DIFFERENCES

William F. Ames, "Numerical Methods for Partial Differential Equations," Barnes and Noble, 1969.

A. R. Mitchell, "Computational Methods in Partial Differential Equations," John Wiley and Sons, 1969.

Gordon D. Smith, "Numerical Solution of Partial Differential Equations," Oxford University Press, 1975.

FINITE ELEMENTS

C. S. Desai and J. F. Abel, "Introduction to the Finite Element Method," van Nostrand-Reinhold, 1972.

A. R. Mitchell and R. Wait, "The Finite Element Method in Partial Differential Equations," John Wiley and Sons, Ltd., 1976.

J. T. Oden, "Finite Elements of Nonlinear Continua," McGraw-Hill Book Co., 1972.

P. M. Prenter, "Splines and Variational Methods," John Wiley and Sons, 1975.

G. Strang and G. Fix, "An Analysis of the Finite Element Method," Prentice-Hall, 1973.

O. C. Zienkiewicz, "The Finite Element Method in Engineering Science," McGraw-Hill Book Co., 1971.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 1705

2. GOVT ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

(9) Technical

4. TITLE (and Subtitle): (6) Solution of Partial Differential Equations By Means Of Finite Elements. An Introductory Sketch

5. TYPE OF REPORT & PERIOD COVERED: Summary Report, no specific reporting period

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s): (10) J. Barkley Rosser

8. CONTRACT OR GRANT NUMBER(s): (15) DAAG29-75-C-0024

9. PERFORMING ORGANIZATION NAME AND ADDRESS: Mathematics Research Center, University of Wisconsin, 610 Walnut Street, Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

11. CONTROLLING OFFICE NAME AND ADDRESS: U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, North Carolina 27709

12. REPORT DATE: December 1976

13. NUMBER OF PAGES: 33

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): (12) 36 p.

15. SECURITY CLASS. (of this report): UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

(14) MRC-TSR-1705

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Partial Differential Equations
Finite Element Methods

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An introductory exposition is given covering the solution of some partial differential equations by means of the method of finite elements. Special attention is given to the means of getting numerical approximations to the answer.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

22200

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)